

Министерство образования и науки Украины
Севастопольский национальный технический университет



**«ПРИМЕНЕНИЕ GRETЛ 1.7.1. ПРИ ПОСТРОЕНИИ И АНАЛИЗЕ
РЕГРЕССИОННЫХ МОДЕЛЕЙ С ГЕТЕРОСКЕДАСТИЧНОЙ
СЛУЧАЙНОЙ СОСТАВЛЯЮЩЕЙ»**

МЕТОДИЧЕСКИЕ УКАЗАНИЯ

к выполнению лабораторной работы №3

по дисциплине «Эконометрия»

**для студентов специальности
6.050201 - «Менеджмент организаций»
дневной формы обучения**

Севастополь
2008



УДК 658

«Применение Gretl 1.7.1. при построении и анализе регрессионных моделей с гетероскедастичной случайной составляющей» методические указания к выполнению лабораторной работы №3 по дисциплине «Эконометрия» для студентов специальности 8.050201 – «Менеджмент организаций» дневной формы обучения / Сост. А.В. Цуканов, Т.А. Кокодей. – Севастополь: Изд-во СевНТУ, 2008г. – 24 с.

Целью методических указаний является получение навыков выполнения регрессионного анализа при наличии гетероскедастичности с использованием средств, предоставляемых Gretl 1.7.1.

Методические указания утверждены на заседании кафедры менеджмента и экономико-математических методов, (протокол № 5 от 23.01.2008 г.).

Допущено учебно-методическим центром СевНТУ в качестве методических указаний.

Рецензент:

Фисун С.Н., канд. техн. наук, доцент кафедры «Кибернетика и вычислительная техника».

СОДЕРЖАНИЕ

1. Цель работы.....	4
2. Теоретические сведения	4
3. Описание средств системы Gretl для выполнения регрессионного анализа при наличии гетероскедастичности.....	7
3.1. Пример обнаружения гетероскедастичности в Gretl	7
3.2. Оценивание гетероскедастичной модели с использованием взвешенного метода наименьших квадратов ВМНК (WLS).....	14
4. Порядок выполнения лабораторной работы	23
5. Содержание отчёта о выполнении лабораторной работы.....	23
Библиографический список.....	24

1. ЦЕЛЬ РАБОТЫ

Цель работы заключается в освоении инструментария системы Gretl в области построения и анализа регрессионных моделей с гетероскедастичной случайной составляющей для выявления и последующего применения ранее неизвестных закономерностей в имеющихся данных в процессе подготовки и принятия решений менеджерами компаний.

2. ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ

Для регрессионной модели, формула (1), построенной по фактическим данным типа срез данных (cross-sectional) дисперсия случайных отклонений (ошибок) часто представляет собой переменную величину $\sigma^2_{u_i} \neq const$.

$$y = \tilde{y}_x + u = \alpha_0 + \alpha_1 \cdot x_1 + \dots + \alpha_n x_n + u, \quad (1)$$

где y — фактическое значение результативного признака;

$\tilde{y}_x$ - модельное значение результативного признака;

α_i — параметр регрессионной модели;

x_i - признак-фактор;

u — случайная ошибка.

Данная ситуация представляет собой проблему гетероскедастичности («неодинакового разброса») - нарушения, возникающего при невыполнении одного из классических предположений линейного регрессионного анализа о постоянстве дисперсий случайных отклонений (гомоскедастичности или «одинакового разброса» $\sigma_{u_j} = const$), при этом остальные условия Гаусса-Маркова выполняются:

- математическое ожидание случайной составляющей, $M(u_i) = 0$
- отсутствие автокорреляции остатков (взаимосвязи u_i и u_{i-1}).
- случайный характер остатков – их независимость от y_i и x_i .

Случаи гетероскедастичности и гомоскедастичности показаны на рисунках 1 и 2 соответственно.

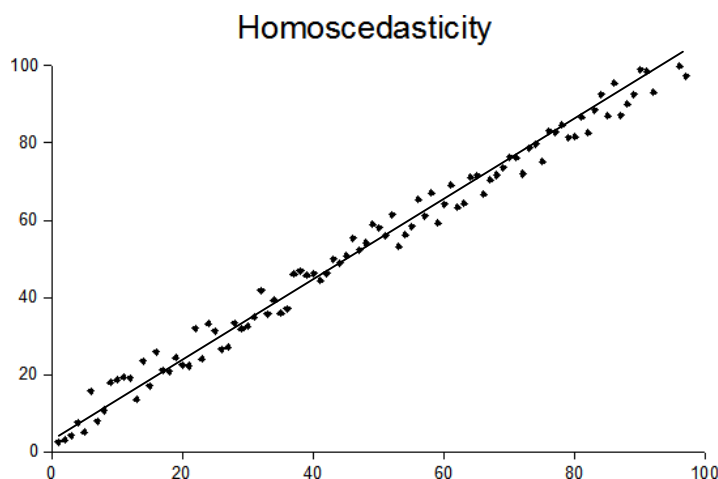


Рисунок 1 - Иллюстрация случайных данных и модели с гомоскедастичностью остатков $\sigma^2_{u_i} = const$

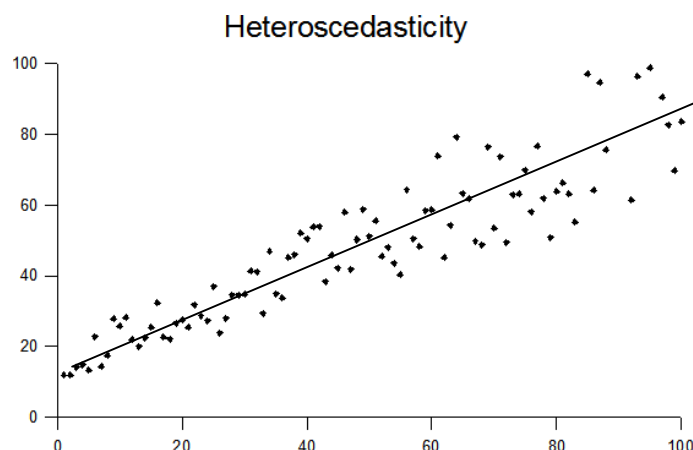


Рисунок 2 - Иллюстрация случайных данных и модели с гетероскедастичностью остатков $\sigma^2_{u_i} \neq const$

Гетероскедастичность приводит к тому, что при применении обычного метода наименьших квадратов (МНК) полученные параметры $\alpha_0 \dots \alpha_n$ модели, формула (1), больше не представляют собой наиболее эффективные оценки или не являются оценками с минимальной дисперсией.

Наблюдение, для которого теоретическое распределение ошибки имеет малое стандартное отклонение, будет обычно находиться близко к линии регрессии и, следовательно, может стать хорошим ориентиром, указывающим на место этой линии. В противоположность этому наблюдение, где теоретическое распределение имеет большое стандартное отклонение, не сможет в той же мере помочь в определении местоположения линии регрессии. Обычный МНК не делает различия между качеством наблюдений, придавая одинаковые "веса" каждому из них независимо от того, является ли наблюдение существенным или несущественным для определения местоположения этой линии. Следовательно, обычным МНК мы получим неэффективные оценки коэффициентов.

Также результаты t- и F- тестов будут ненадёжными, т.к. мы получим неверные оценки стандартных ошибок параметров $\alpha_0 \dots \alpha_i$ (STDERROR), т.к. они вычисляются на основе предположения о том, что остатки модели гомоскедастичны, что скажется на правильности расчёта t- и F- статистик и приведёт к принятию ошибочных гипотез.

Поскольку в данном случае использование обычного метода наименьших квадратов (МНК) неэффективно, необходимо сделать поправку на гетероскедастичность, применив взвешенный метод наименьших квадратов ВМНК (WLS) для её устранения.

Построение гетероскедастичной регрессионной модели состоит из двух этапов:

1. Обнаружение гетероскедастичности случайной составляющей,

2. Оценивание модели с использованием взвешенного метода наименьших квадратов (WLS).

На первом этапе в случае однофакторной регрессии $y_i = f(x_i) + u_i = \alpha_0 + \alpha_1 \cdot x_i + u_i$ изначально проводится графический анализ остатков – строится и анализируется зависимость квадратов ошибок от x_i или от теоретического значения $\tilde{y}_{x_i}$, или строится $x_i - y_i$ диаграмма рассеяния. При множественной регрессии графический анализ также возможен для каждой из объясняющих переменных. Рост дисперсии с ростом одного из факторов свидетельствует о гетероскедастичности.

Затем проводится один из формальных тестов на гетероскедастичность: тест ранговой корреляции Спирмена, тест Парка (The Park test), тест Голдфелда-Квандта (Goldfeld-Quandt test), тест Бреуша-Погана, тест Глейзера, или тест Уайта (White's test) и осуществляется интерпретация результатов теста.

В каждом тесте пытаются опровергнуть гипотезу о гомоскедастичности, если это удаётся, то можно сделать вывод, что в модели наблюдается гетероскедастичность.

Рассмотрим алгоритм теста Уайта на гетероскедастичность, не требующего нормальности распределения остатков. Алгоритм состоит из следующих шагов:

- Получение остатков оцененной регрессионной модели $u = y - \alpha_0 - \alpha_1 \cdot x_1 - \dots - \alpha_n x_n$
- Оценивание вспомогательного уравнения регрессии квадратов остатков относительно комплекса переменных модели, их произведений и их квадратов $u^2 = \beta_0 + \beta_1 \cdot x_1 + \dots + \beta_n x_n + \beta_{n+1} \cdot x_1^2 + \dots + \beta_{2n} x_n^2 + \beta_{2n+1} \cdot x_1 \cdot x_2 + \dots + \beta_p \cdot x_n \cdot x_{n-1} + \varepsilon$
- Проверка общей значимости уравнения с помощью критерия χ^2 . Тестовой статистикой является величина $n \cdot R^2$ (n - число наблюдений, R^2 - коэффициент детерминации). Число степеней свободы равно числу регрессоров вспомогательного уравнения. Если $n \cdot R^2 \succ \chi_{\text{крит}}^2$, то нулевая гипотеза гомоскедастичности $H_0: \beta_0 = \beta_1 = \dots = \beta_i = 0$ отвергается.

На втором этапе для оценки моделей с гетероскедастичностью используется взвешенный метод наименьших квадратов (ВМНК - WLS). Метод ВМНК, как и МНК, применим к однофакторной и множественной линейной регрессии и использует то же правило минимизации суммы квадратов остатков, RSS, но вместо одинаковых весов для каждого наблюдения им приписываются значения, обратные соответствующим дисперсиям ошибки, что отражено формулой (2).

$$\sum_{i=1}^n w_i (y_i - \tilde{y}_{x_i})^2 = \sum_{i=1}^n \frac{1}{\sigma_{u_i}^2} (y_i - \tilde{y}_{x_i})^2 \rightarrow \min, \quad (2)$$

где y_i, y_i - фактическое и модельное i -е значение зависимой переменной;

w_i – вес i -го наблюдения, $w_i = \frac{1}{\sigma_{u_i}^2}$;

$\sigma_{u_i}^2$ - дисперсия i -й случайной составляющей.

Тогда коэффициенты линейной регрессии находятся по формуле (3).

$$\mathcal{E} = [X^T \cdot W \cdot X]^{-1} \cdot X^T \cdot W \cdot Y, \quad (3)$$

где $W = \text{diag}\{w_1, \dots, w_n\}$ - диагональная матрица весов;
 n - число наблюдений.

Дисперсия ошибки $\sigma_{u_i}^2$ чаще всего неизвестна, но возможно существование некоторого соотношения между дисперсией ошибки и значением какой-либо объясняющей переменной в регрессионной модели $y_i = \alpha_0 + \alpha_1 x_{1i} + \dots + \alpha_p x_{pi} + u_i$, например, $\sigma_{u_i}^2 = c \cdot x_{1i}^2$, где c - ненулевая константа и x_{1i} - значение объясняющей переменной x_1 в i -ом наблюдении. В случае подобного соотношения $\sigma_{u_i}^2$ можно считать известными, т.к. постоянная величина «с» не влияет на взвешенную процедуру. Тогда значения весов $w_i = \frac{1}{c \cdot x_{1i}^2}$.

3. ОПИСАНИЕ СРЕДСТВ СИСТЕМЫ GRETL ДЛЯ ВЫПОЛНЕНИЯ РЕГРЕССИОННОГО АНАЛИЗА ПРИ НАЛИЧИИ ГЕТЕРОСКЕДАСТИЧНОСТИ

3.1. Пример обнаружения гетероскедастичности в Gretl

Шаг 1. Выберем команду меню **File\Open Data\Sample File...** и на закладке **Verbeek** двойным щелчком мыши откроем встроенный разработчиками набор данных **bwages.gdt** (рисунок 3).

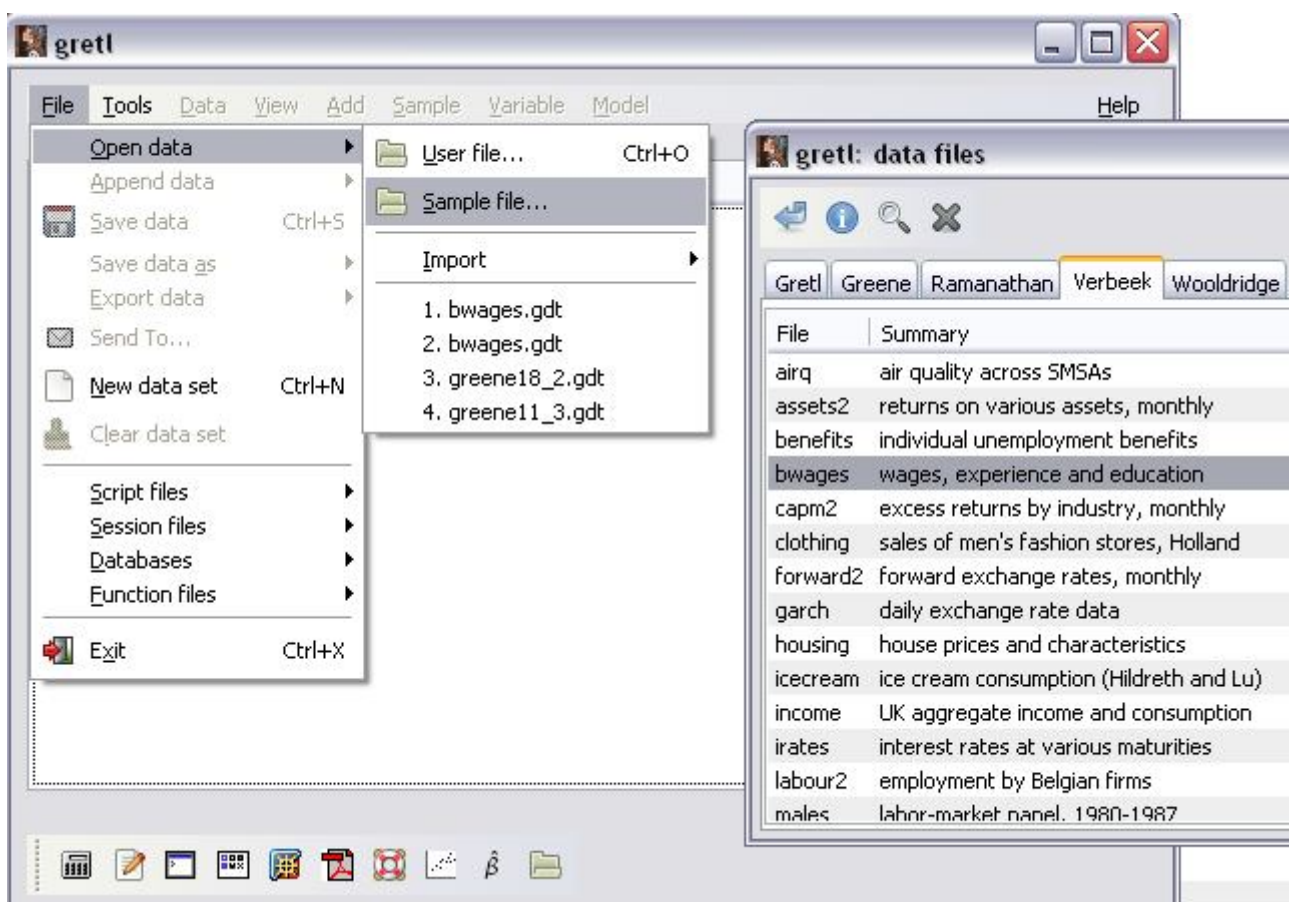


Рисунок 3 – Открытие встроенного набора данных bwages.gdt

Шаг 2. Просмотрим текстовую информацию о наборе данных, выбрав команду **Data\Print Description**. В открывшемся окне (рисунок 4) появится информация о наборе данных **bwages.gdt** в целом и о каждой переменной.

Файл содержит 1472 наблюдения, относящихся к 1994, группе бельгийских семей Европейского Сообщества. Тип данных **undated** (срез данных для фиксированного момента времени – cross-sectional). Переменные:
wage – заработная плата в час, до выплаты налогов (Евро),
educ – уровень образования от 1 (низкий) до 5 (высокий),
exper – опыт работы (лет)
male – фиктивная переменная, принимает значение «1», если мужчина и «0», если женщина. Также представлены логарифмы переменных wage, educ, exper. Требуется оценить влияние уровня образования (educ) в бельгийских семьях на величину заработной платы (wage), используя фактические данные 1994 года по 1472 семьям.

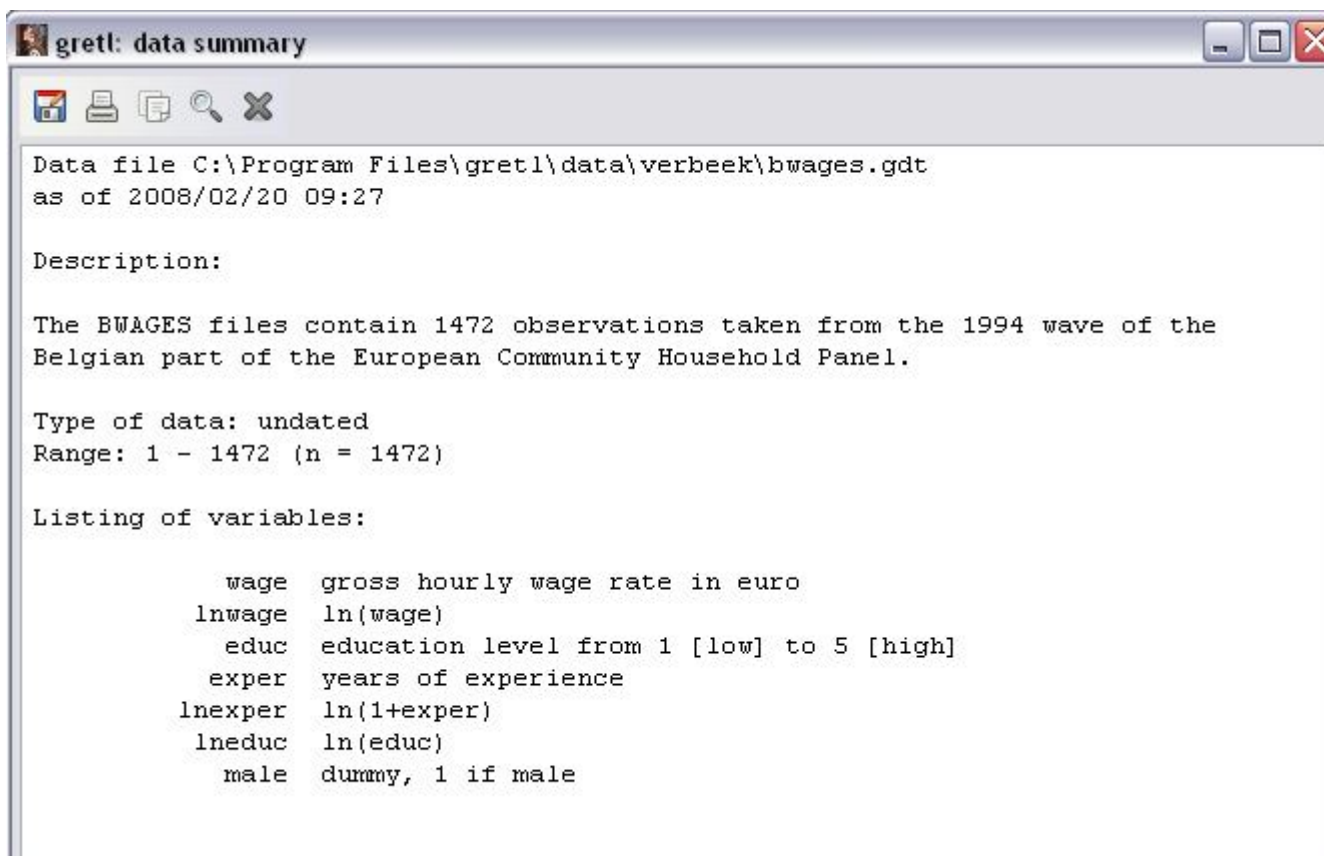


Рисунок 4 – Текстовое описание набора данных bwages.gdt

Шаг 3. Проведём предварительный графический анализ данных, построив диаграммы рассеяния пар переменных educ -wage, exper-wage, male – wage.

Выберем команду меню **View\Multiple graphs Vars\ X-Y scatters...**(рисунок 5). В открывшемся окне выберем переменную wage (по оси OY), а переменные educ, exper, male (по оси OX) нажатием кнопок «Choose» и «Add» соответственно.

Как показывает рисунок 6, дисперсия wage (которую будем рассматривать как зависимую переменную Y) увеличивается при росте каждого из факторов **educ**, **exper**, **male** (которые будем рассматривать как независимые переменные, X_i), что свидетельствует о гетероскедастичности.

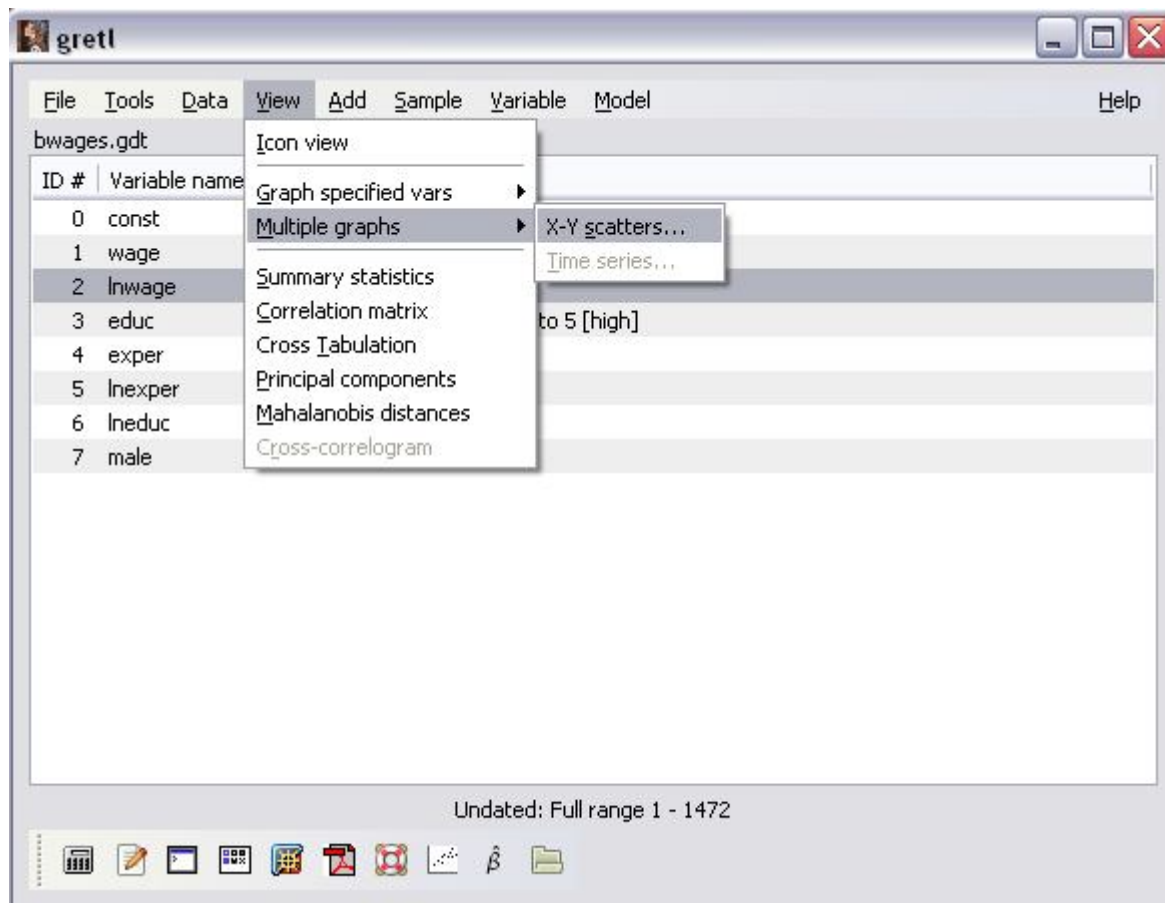


Рисунок 5 – Построение диаграмм рассеяния educ-wage, exper-wage, male-wage

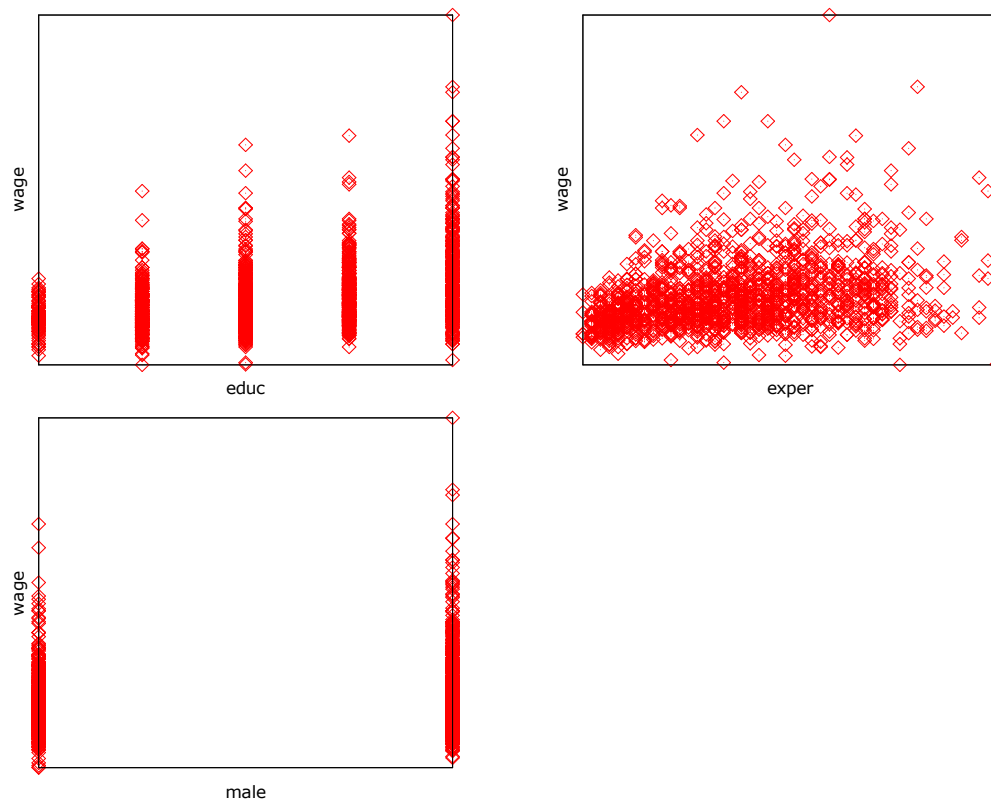


Рисунок 6 – Диаграммы рассеяния пар переменных educ -wage, exper-wage, male -wage

Шаг 4. Проведём оценивание модели $wage = \alpha_0 + \alpha_1 \cdot educ + u$ обычным методом наименьших квадратов 1МНК (OLS). Для этого выберем команду меню **Model\Ordinary Least Squares...** и в открывшемся окне спецификации модели выберем зависимую переменную wage при помощи кнопки «Choose» и независимую переменную educ при помощи кнопки «Add». После нажатия кнопки ОК появится окно с результатами моделирования (рисунок 7).

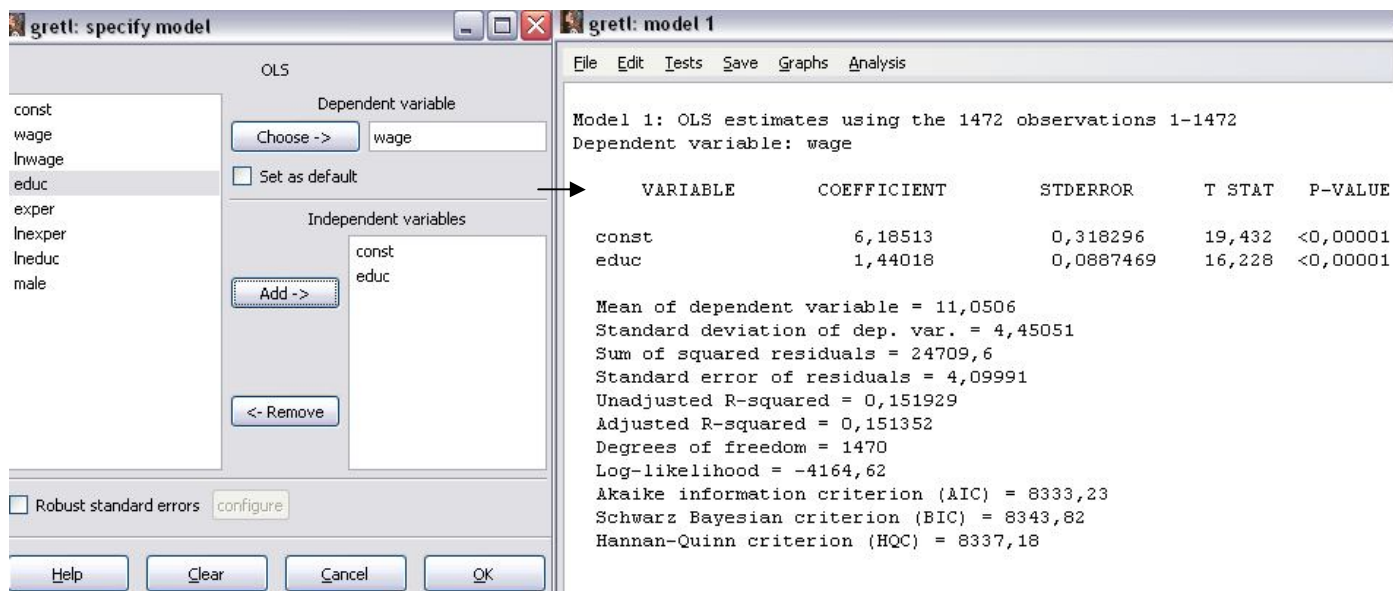


Рисунок 7 – Построение модели $wage = \alpha_0 + \alpha_1 \cdot educ + u$ методом 1МНК (OLS)

По результатам 1МНК в полученной модели $wage = 6,18513 + 1,44018 \cdot educ$ параметры α_0, α_1 являются значимыми при уровнях значимости 5% и 1%, поскольку $P\text{-VALUE} = 0,001\% < 1\% < 5\%$ (принимая альтернативную гипотезу $H_1: \alpha_1 \neq 0$).

Хотя по данным 1МНК можно установить существенность влияния переменной educ (образование) на переменную wage (з.пл.), необходимо сделать поправку на гетероскедастичность в связи с ошибочностью расчёта величин COEFFICIENT, STDERROR и T-STAT и ненадёжностью результатов данного оценивания -возможностью принятия неправильной гипотезы.

Сохраним величины квадратов остатков в отдельную переменную usq1 набора данных, обратившись к команде **Save\Squared Residuals** меню окна результатов моделирования (рисунок 7) и нажав кнопку ОК.

Шаг 5. Подтвердим наличие гетероскедастичности, используя формальный тест Уайта. Обратимся к команде **Tests\Heteroskedasticity** меню окна результатов моделирования (рисунок 8).

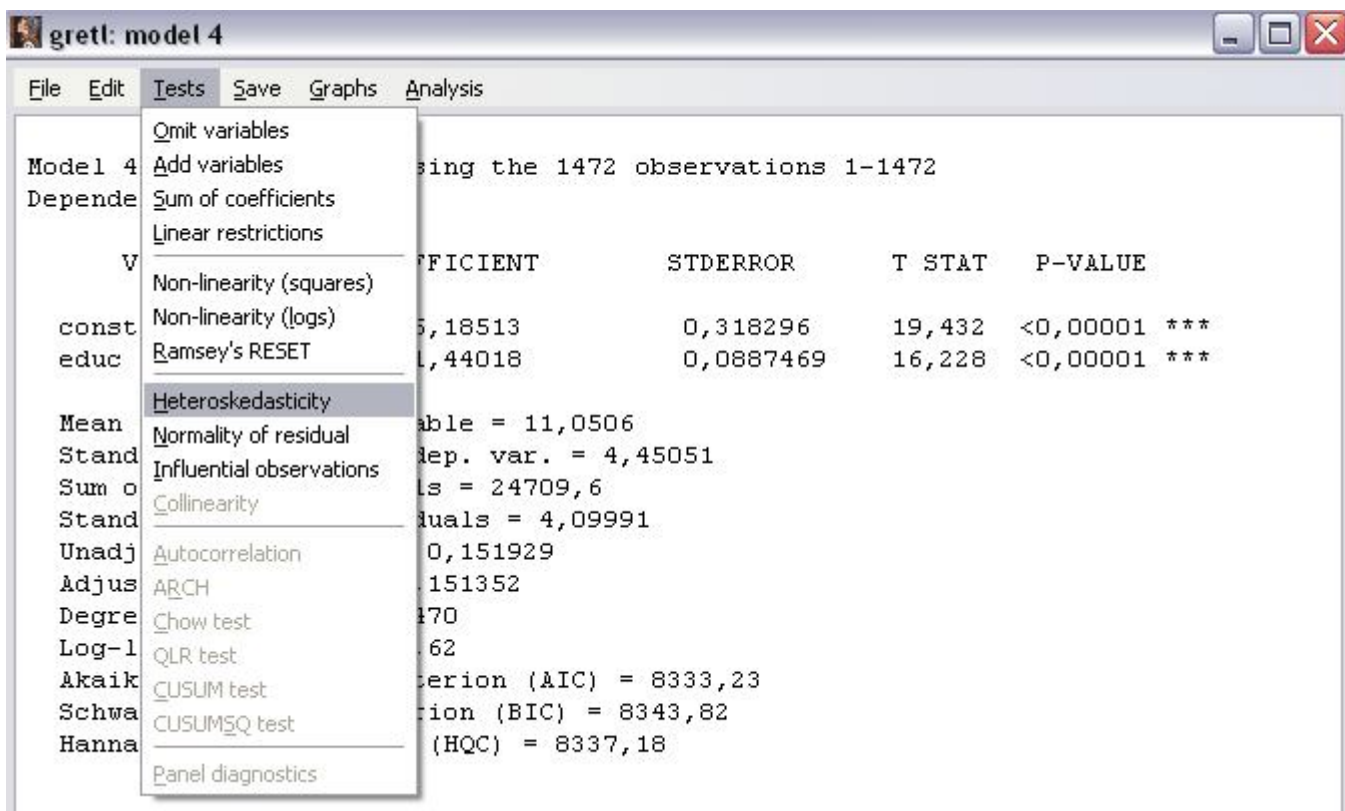


Рисунок 8 – Проведение теста Уайта на гетероскедастичность

В результате теста получим вспомогательную модель регрессии остатков относительно переменной *educ* и её квадрата $u^2 = 16,1382 - 10,3147 \cdot educ + 2,7594 \cdot educ^2$ (рисунок 9).

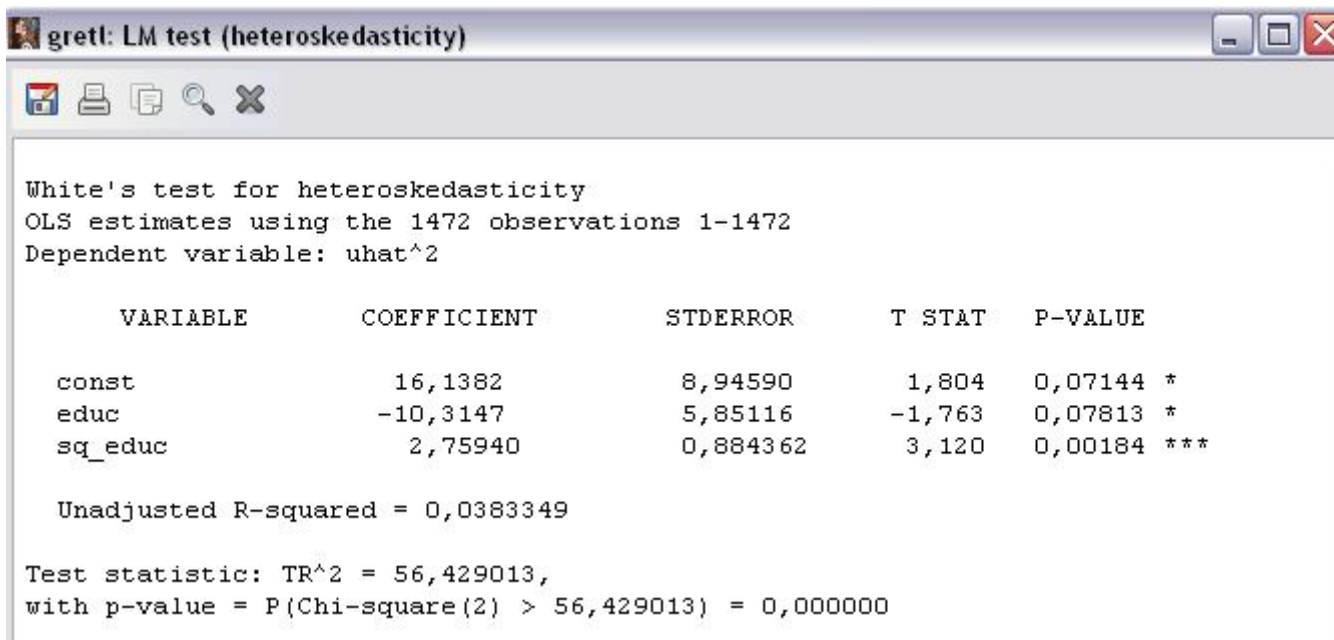


Рисунок 9 – Окно результатов теста Уайта на гетероскедастичность

Проверим общую значимость (адекватность в целом) данной модели, используя критерий χ^2 .

В окне результатов теста (рисунок 9) значение p-value для статистики теста $n \cdot R^2$ (TR^2) составило 0,000%, что меньше уровней значимости 1% и 5% и свидетельствует о наличии гетероскедастичности (адекватности вспомогательной модели, неравенстве нулю всех её параметров).

Т.о. отвергаем гипотезу H_0 о гомоскедастичности (равенстве нулю всех её параметров), поскольку расчётное значение статистики $n \cdot R^2 = TR^2 = 56,429$ больше χ^2 критического ($n \cdot R^2 > \chi_{crit}^2$).

Найдём χ^2 критическое для уровня значимости 1% , обратившись к команде основного меню **Tools\Statistical Tables** и введя в открывшемся окне на закладке **Chi-Square** число степеней свободы (2, равное числу регрессоров ($educ, educ^2$) вспомогательной модели) и уровень значимости 0.001 (рисунок 10). Получим $\chi_{crit}^2 = 13,8155$ (Critical value).

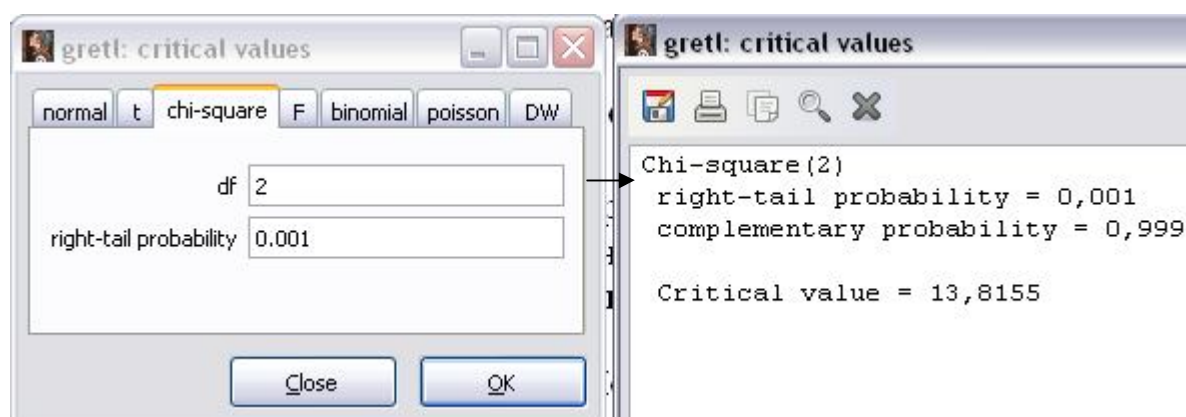


Рисунок 10 – Нахождение критического значения χ^2 для $df=2$, $p=1\%$

Построим график данного χ^2 распределения, выбрав команду основного меню **Tools\Distribution Graphs**, закладку Chi-square и $df=2$ в открывшемся окне (рисунок 11).

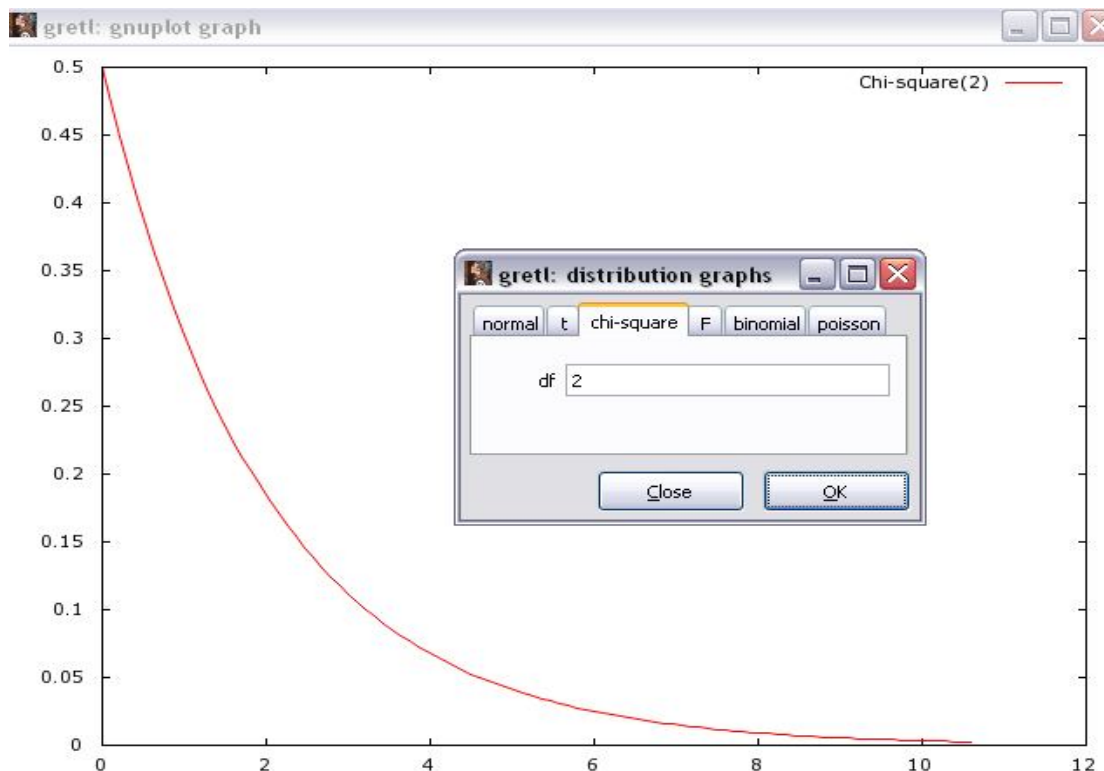


Рисунок 11 - График χ^2 распределения, df=2

3.2. Оценивание гетероскедастичной модели с использованием взвешенного метода наименьших квадратов ВМНК (WLS)

Существует два способа реализации ВМНК (WLS) в пакете Gretl:

1. При выборе из главного меню команды **Model\Other Linear Models\Weighted Least Squares....**(рисунок 12) открывается окно спецификации модели, которое предусматривает выбор из списка переменных открытого набора переменных - веса w . Данная переменная добавляется к набору путём ввода данных вручную или определяется путём ввода соответствующей формулы. Данный способ используется если дисперсия ошибки $\sigma_{u_i}^2$ или известна или неизвестна, но существует соотношение между ней и одной из объясняющих переменных (x_i), например, $\sigma_{u_i}^2 = c \cdot x_{ii}^2$.
2. При выборе из главного меню команды **Model\Other Linear Models\Heteroskedasticity Corrected...** (рисунок 12) веса наблюдений (w_i) определяются Gretl по формуле (4) и не вводятся пользователем. В качестве условия корректности применения рассматриваемого метода оценивания рассматривается получение нормального распределения остатков. Форму распределения можно проверить при помощи команды меню окна результатов моделирования **Tests\normality of residual....**

$$w_i = \frac{1}{\sqrt{e^{\hat{y}_i}}}, \quad (4)$$

где $\tilde{y}_i$ - модельные значения зависимой переменной, полученные при использовании 1МНК; $e=2,718..$; w_i -вес i-го наблюдения.

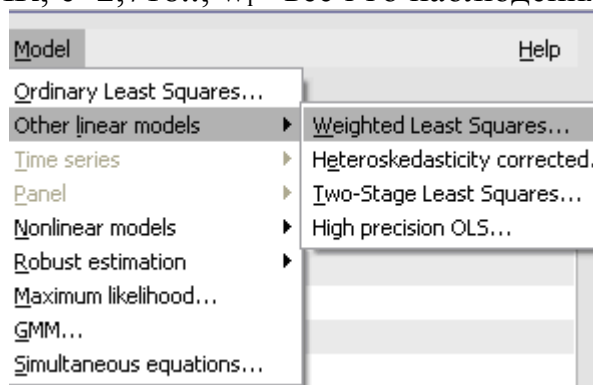


Рисунок 12 – Способы реализации ВМНК (WLS) в пакете Gretl

Обратимся к вышеприведённому примеру набора данных bwages.gdt для иллюстрации обоих способов оценивания регрессионной модели $wage = \alpha_0 + \alpha_1 \cdot educ + u$ методом ВМНК.

Этап 1. Введём новую переменную – вес w , которая будет определяться по формуле $w_i = \frac{1}{\sigma_{u_i}^2} = \frac{1}{3 \cdot educ_i^2}$.

Для этого выберем команду Define new variable из меню Variable (рисунок 13).

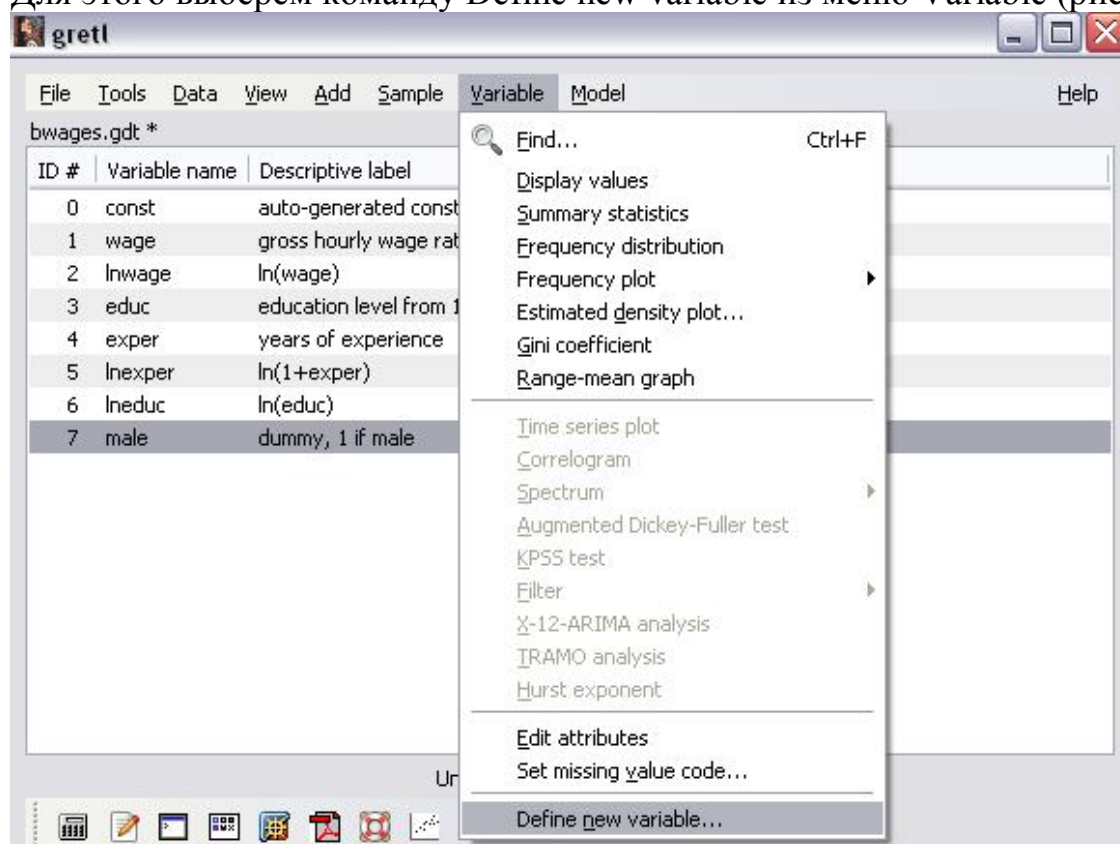


Рисунок 13 – Создание новой переменной

В открывшемся окне “add var” введём имя переменной *w* и нажмём ОК, в появившемся окне «edit data» не вводим значения переменной, а нажимаем кнопку «close» для создания переменной с пустыми значениями (рисунок 14).

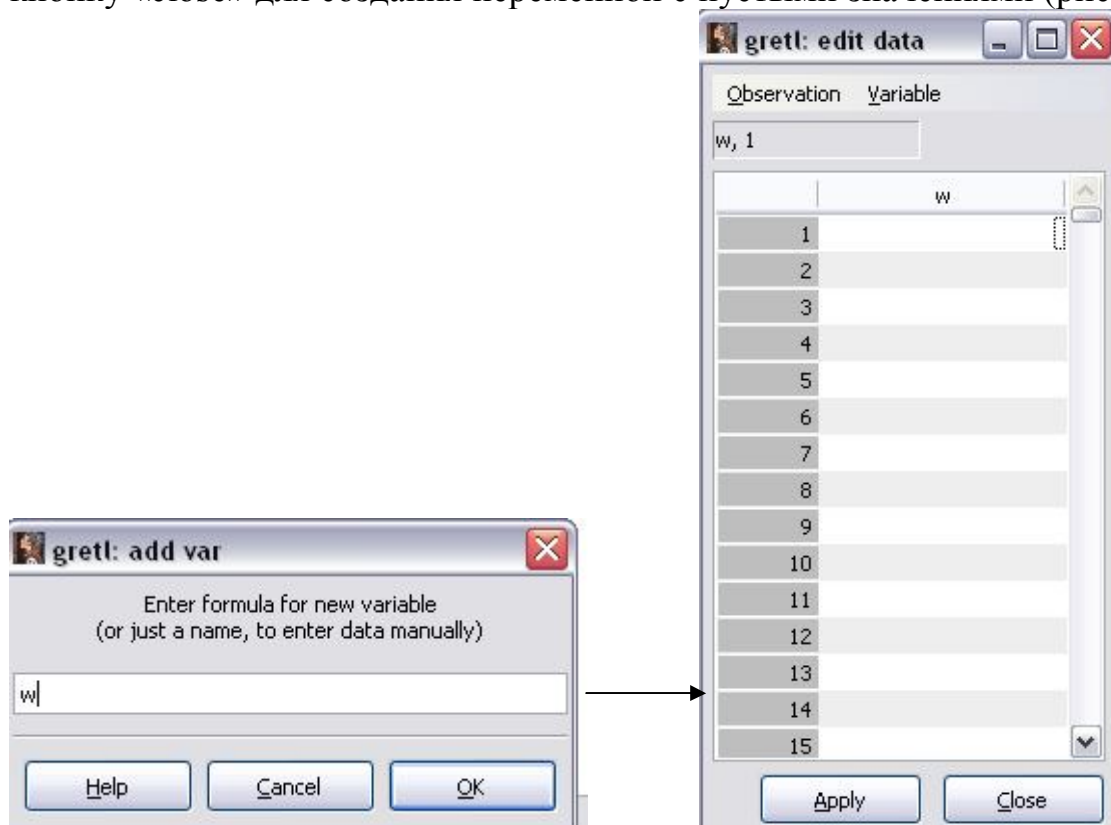


Рисунок 14 – Ввод имени переменной

Данная переменная *w* с пустыми значениями и без описания появится в списке открытого набора данных. Введём описание переменной, обратившись к команде **Variable\Edit Attributes**.

В открывшемся окне (рисунок 15) введём в поле “Description” описание «weights» (веса) и нажмём ОК.

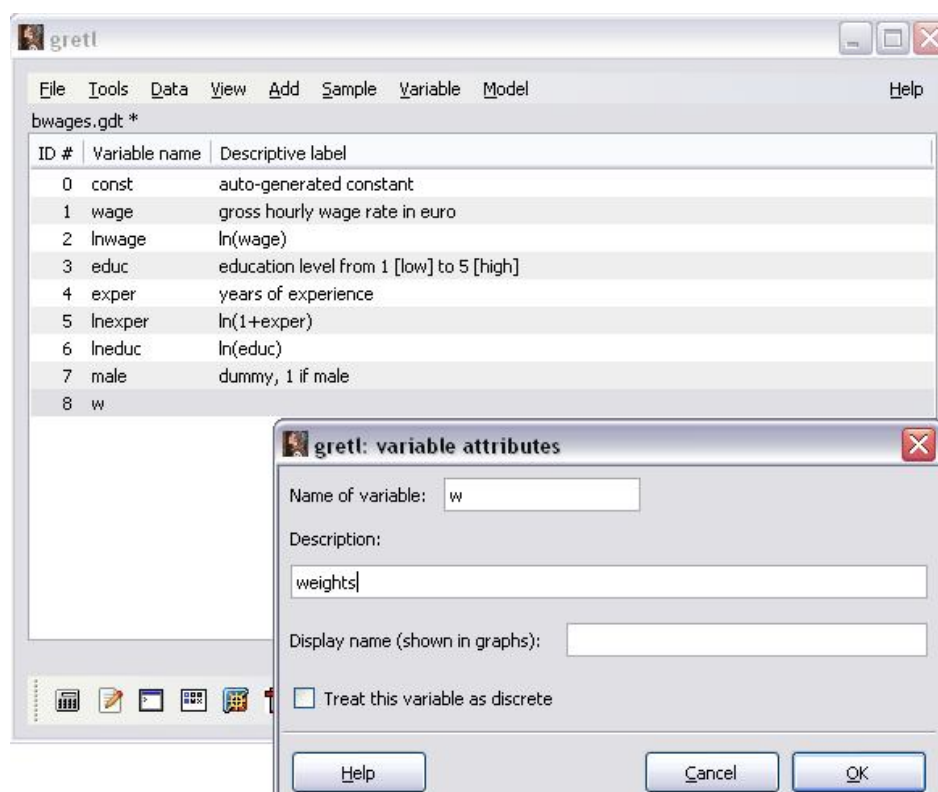
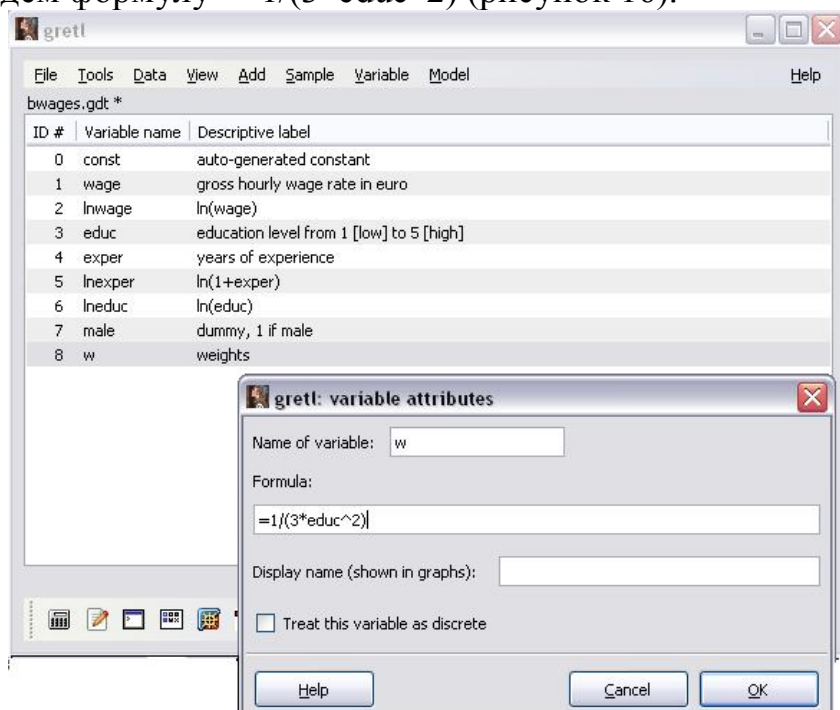
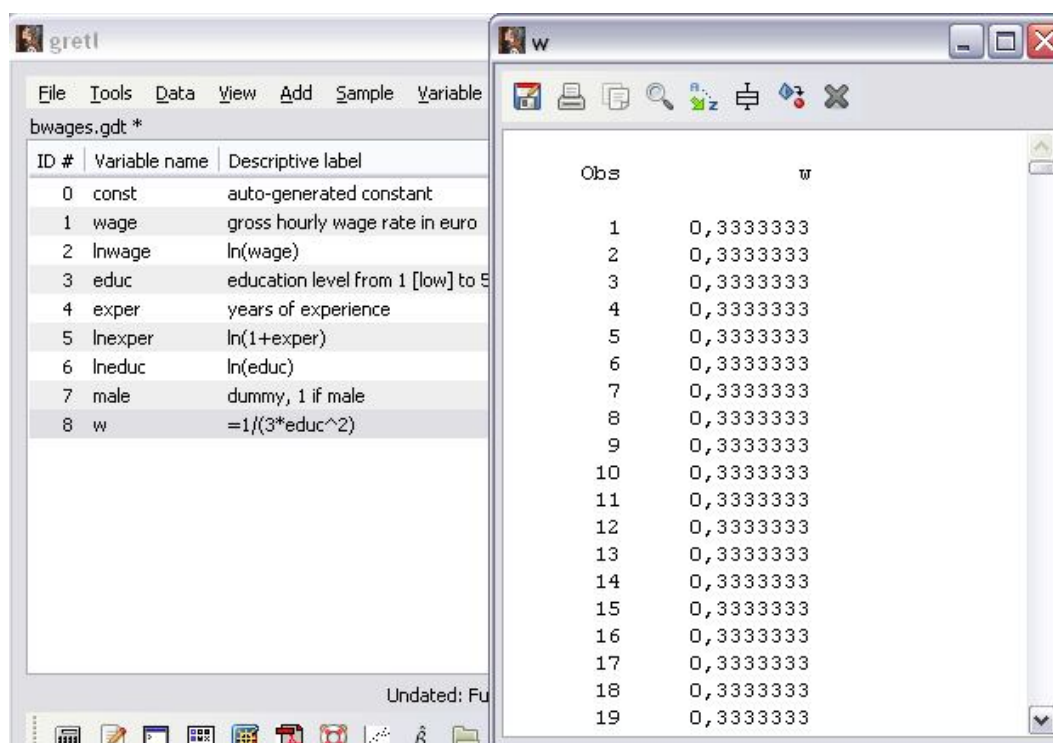


Рисунок 15 – Изменение атрибутов переменной

Введём формулу для переменной w , повторно обратившись к команде **Variable\Edit Attributes**. Теперь в поле “Description” вместо описания «weights» введём формулу $=1/(3*\text{educ}^2)$ (рисунок 16).

Рисунок 16 – Ввод формулы для определения переменной w

Этап 2. Просмотрим значения созданной переменной w , выбрав её в списке щелчком мыши и обратившись к команде **Data\Display Values** (или двойным щелчком мыши по её названию в списке) (рисунок 17).



The screenshot shows the Gretl interface with two windows. The main window displays a list of variables with their IDs, names, and descriptive labels. The variable 'w' is highlighted, with its formula $w = 1/(3 * educ^2)$ shown. A secondary window titled 'w' displays the values of this variable for 19 observations, all of which are 0.3333333.

ID #	Variable name	Descriptive label
0	const	auto-generated constant
1	wage	gross hourly wage rate in euro
2	lnwage	ln(wage)
3	educ	education level from 1 [low] to 5
4	exper	years of experience
5	lnexper	ln(1+exper)
6	lneduc	ln(educ)
7	male	dummy, 1 if male
8	w	$=1/(3 * educ^2)$

Obs	w
1	0,3333333
2	0,3333333
3	0,3333333
4	0,3333333
5	0,3333333
6	0,3333333
7	0,3333333
8	0,3333333
9	0,3333333
10	0,3333333
11	0,3333333
12	0,3333333
13	0,3333333
14	0,3333333
15	0,3333333
16	0,3333333
17	0,3333333
18	0,3333333
19	0,3333333

Рисунок 17 – Просмотр значений переменной w

Этап 3. Для оценки модели $wage = \alpha_0 + \alpha_1 \cdot educ + u$ методом ВМНК (OLS) выполним команду **Model\Other Linear Models\Weighted Least Squares....**(рисунок 18) и в открывшемся окне спецификации модели (рисунок 19) выберем при помощи кнопок «Choose» и «Add» переменные: wage (з.пл) - зависимая переменная; w – веса наблюдений; educ (уровень образования) – независимая переменная.

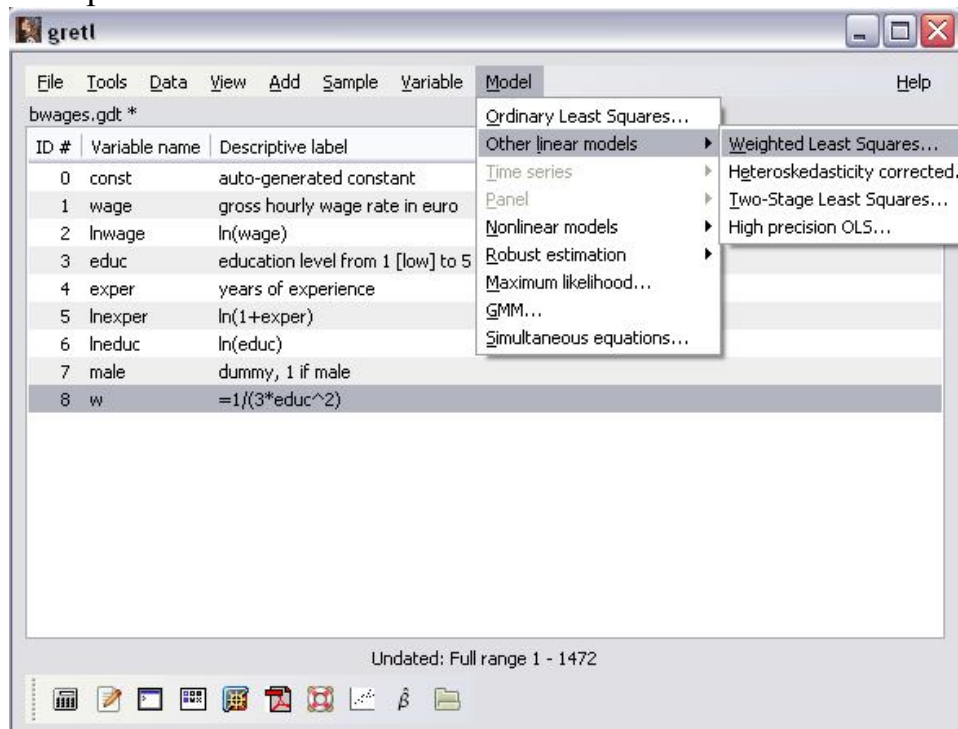


Рисунок 18 – Реализация ВМНК (OLS) в Gretl



Рисунок 19 – Оценивание модели $wage = \alpha_0 + \alpha_1 \cdot educ + u$ при помощи ВМНК (OLS)

Результаты оценивания представлены в окне результатов моделирования (рисунок 20).

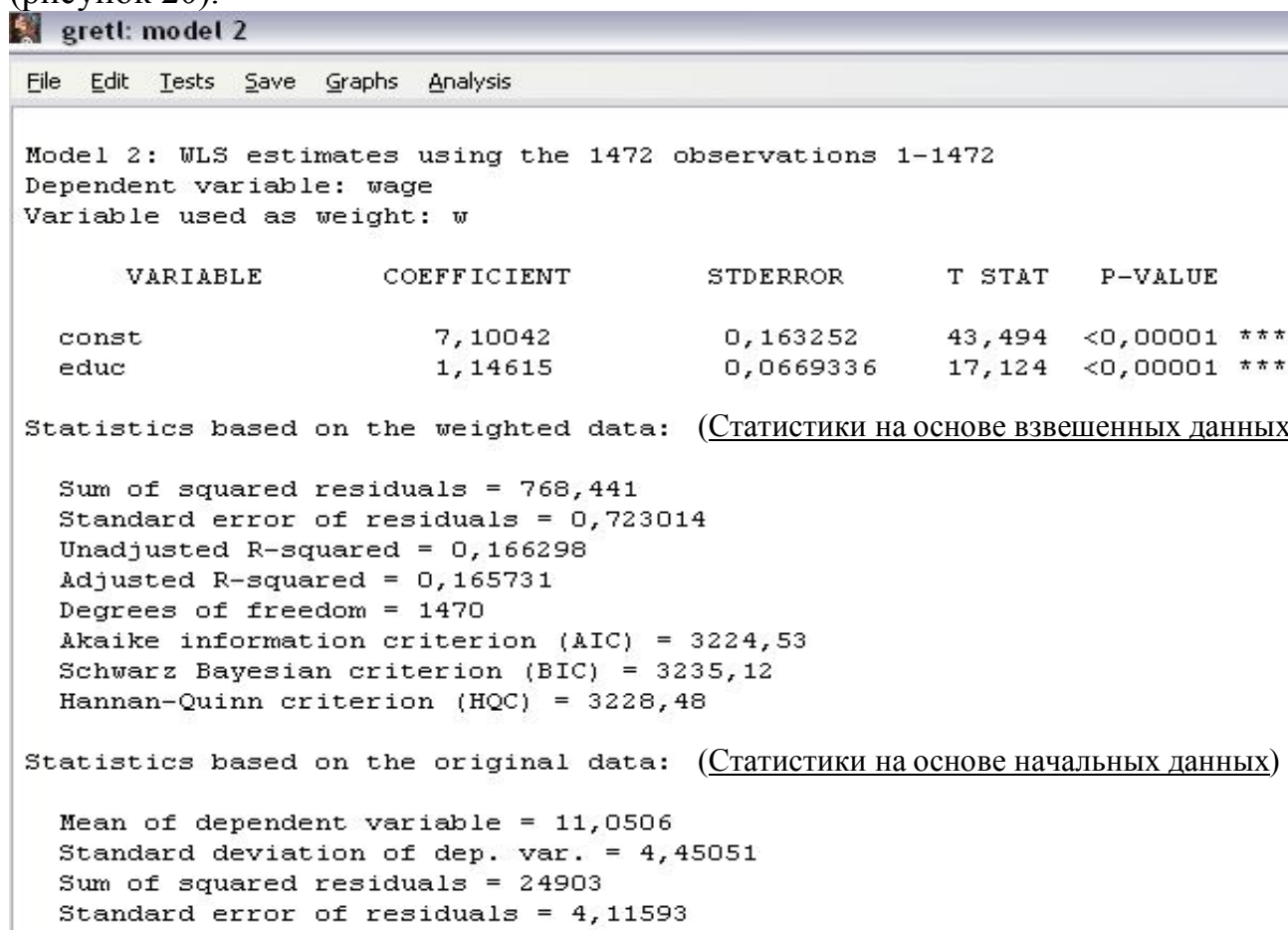


Рисунок 20 – Результаты оценивания модели $wage = \alpha_0 + \alpha_1 \cdot educ + u$ методом ВМНК (WLS)

Сравнивая модели, полученные методом ВМНК $wage = 7,10042 + 1,14615 \cdot educ$ (рисунок 20) и методом 1МНК $wage = 6,18513 + 1,44018 \cdot educ$ (рисунок 7) можно сделать вывод о том, что при использовании 1МНК оценка коэффициента α_1 была завышена, а оценки α_0 и T-STAT – занижены, но на правильности решения о принятии альтернативной гипотезы это не отразилось (значения P-VALUE не изменились).

Этап 4. Воспользуемся вторым способом оценки модели $wage = \alpha_0 + \alpha_1 \cdot educ + u$ методом ВМНК (OLS), выполнив команду **Model\Other Linear Models\Heteroskedasticity Corrected...** (рисунок 21).

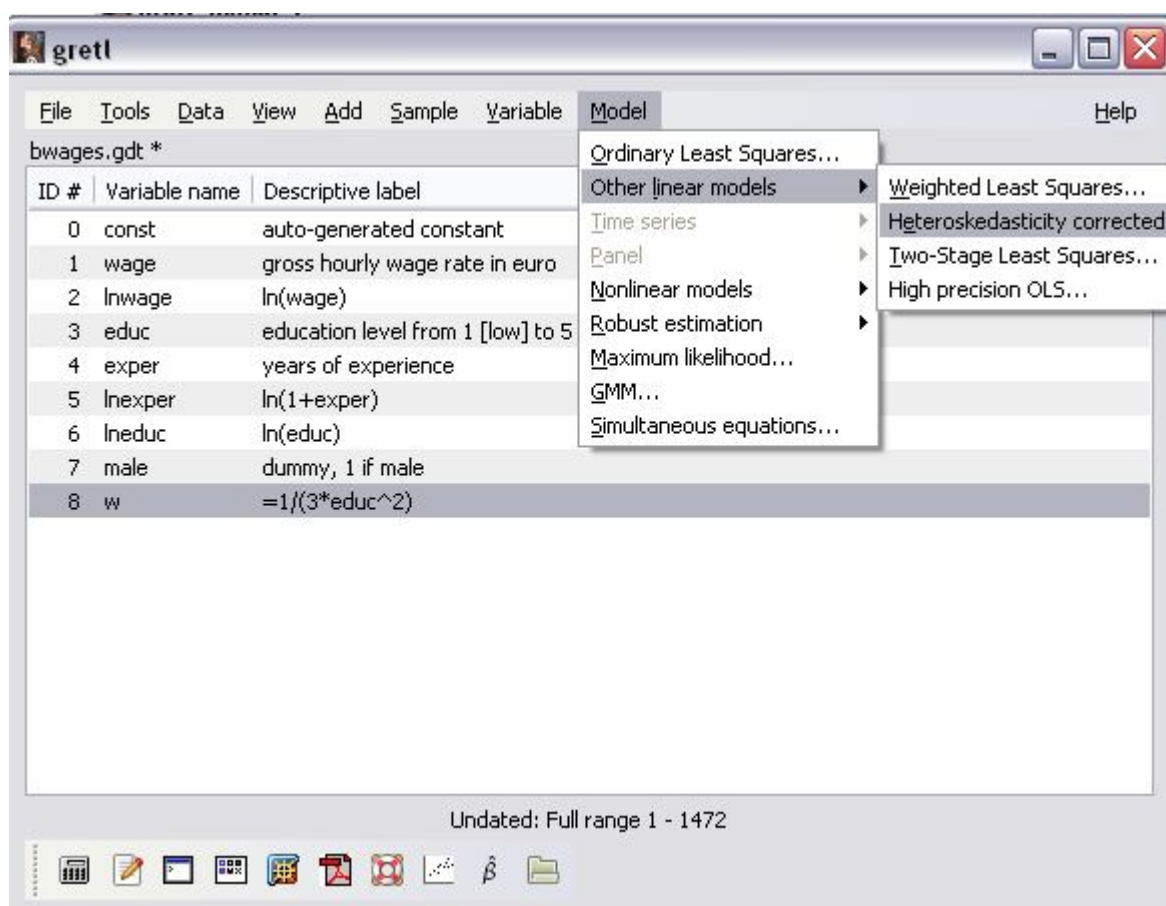


Рисунок 21 – Альтернативный способ оценки модели методом ВМНК (WLS)

Заполним окно спецификации также, как на предыдущем этапе за исключением ввода переменной – веса w , и нажмём кнопку ОК (рисунок 22).

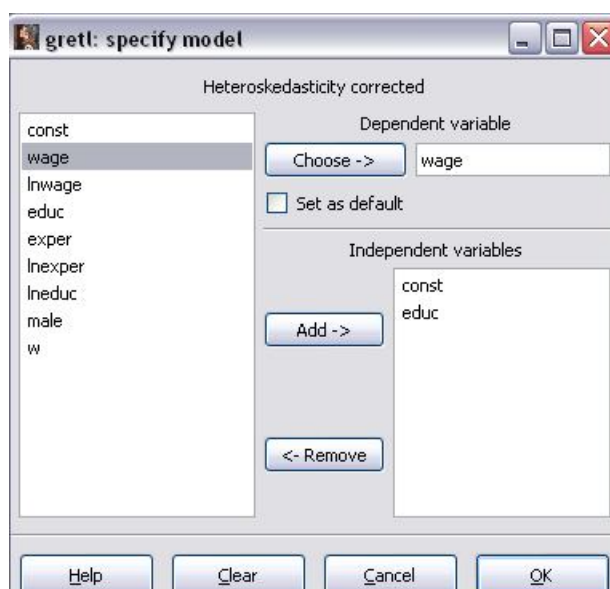


Рисунок 22 – Оценивание модели $wage = \alpha_0 + \alpha_1 \cdot educ + u$ при помощи ВМНК (OLS) с автоматическим определением весов наблюдений

Сравнивая полученную модель $wage = 6,74951 + 1,2509 \cdot educ$ (рисунок 23) с моделью, оцененную с применением 1МНК $wage = 6,18513 + 1,44018 \cdot educ$ (рисунок 7) можно отметить расхождения в значениях показателей регрессионной таблицы, но на правильности решения о принятии альтернативной гипотезы это также не отразилось (значения P-VALUE не изменились).

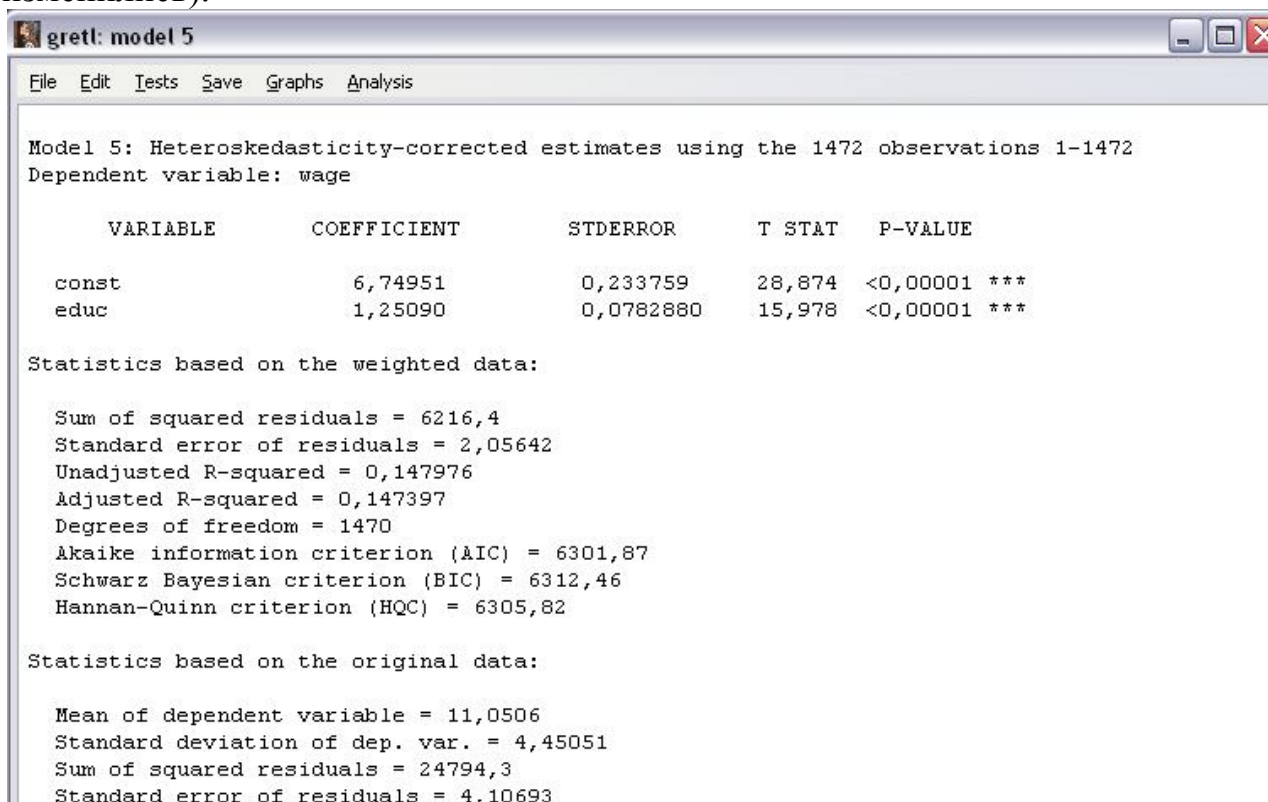
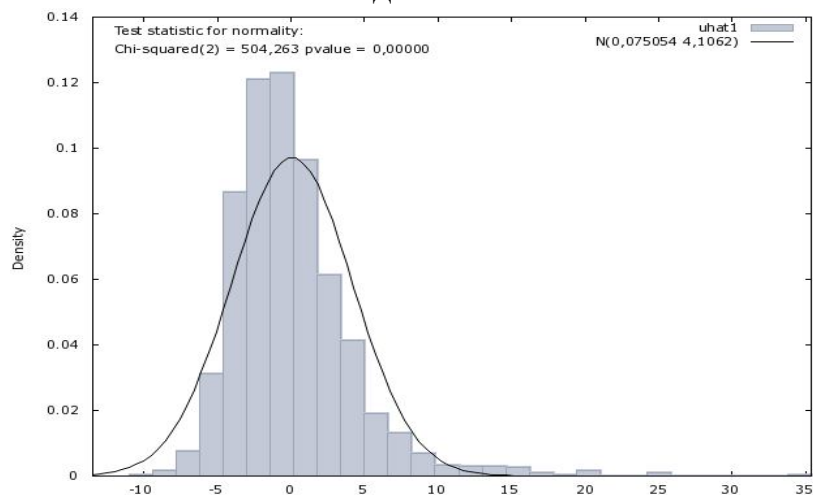


Рисунок 23 – Результаты оценивания модели $wage = \alpha_0 + \alpha_1 \cdot educ + u$ методом ВМНК (WLS) с автоматическим определением весов наблюдений

Проверим условие корректности применения рассматриваемого метода – получение нормального распределения остатков, - обратившись к команде **Tests\normality of residual...** меню окна модели (рисунок 23).

Рисунок 24 показывает, что данное условие не соблюдается $P\text{-value}=0,0000$, что меньше уровней значимости 1% и 5%, поэтому отвергаем нулевую гипотезу о нормальности распределения остатков.

Вывод: В результате исследования 1479 бельгийских семей выявлена прямая зависимость заработной платы (*wage*) от уровня образования (*educ*). Наилучшей моделью, описывающей рассматриваемую зависимость, является модель $wage = 7,10042 + 1,14615 \cdot educ$, полученная методом ВМНК (WLS) с заданием пользователем весов наблюдений.



Frequency distribution for uhat1, obs 1-1472
number of bins = 29, mean = 0,0750541, sd = 4,10624

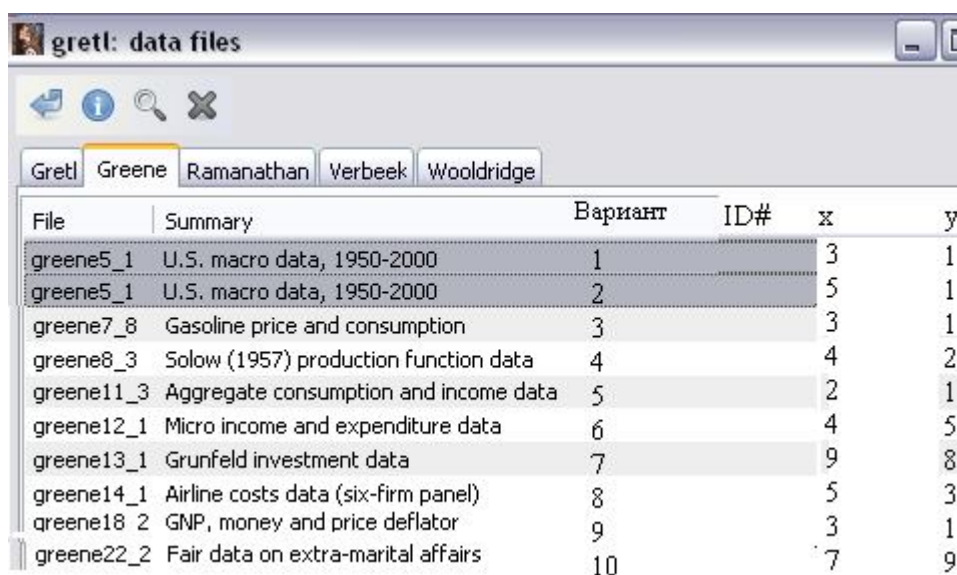
interval	midpt	frequency	rel.	cum.
< -9,3878	-10,187	1	0,07%	0,07%
-9,3878 - -7,7893	-8,5885	4	0,27%	0,34%
-7,7893 - -6,1908	-6,9900	18	1,22%	1,56%
-6,1908 - -4,5922	-5,3915	73	4,96%	6,52%
-4,5922 - -2,9937	-3,7930	204	13,86%	20,38%
-2,9937 - -1,3952	-2,1945	285	19,36%	39,74%
-1,3952 - 0,20332	-0,59594	289	19,63%	59,38%
0,20332 - 1,8018	1,0026	227	15,42%	74,80%
1,8018 - 3,4004	2,6011	144	9,78%	84,58%
3,4004 - 4,9989	4,1996	97	6,59%	91,17%
4,9989 - 6,5974	5,7981	45	3,06%	94,23%
6,5974 - 8,1959	7,3967	31	2,11%	96,33%
8,1959 - 9,7944	8,9952	16	1,09%	97,42%
9,7944 - 11,393	10,594	8	0,54%	97,96%
11,393 - 12,991	12,192	7	0,48%	98,44%
12,991 - 14,590	13,791	7	0,48%	98,91%
14,590 - 16,189	15,389	6	0,41%	99,32%
16,189 - 17,787	16,988	2	0,14%	99,46%
17,787 - 19,386	18,586	1	0,07%	99,52%
19,386 - 20,984	20,185	4	0,27%	99,80%
20,984 - 22,583	21,783	0	0,00%	99,80%
22,583 - 24,181	23,382	0	0,00%	99,80%
24,181 - 25,780	24,980	2	0,14%	99,93%
25,780 - 27,378	26,579	0	0,00%	99,93%
27,378 - 28,977	28,177	0	0,00%	99,93%
28,977 - 30,575	29,776	0	0,00%	99,93%
30,575 - 32,174	31,374	0	0,00%	99,93%
32,174 - 33,772	32,973	0	0,00%	99,93%
>= 33,772	34,571	1	0,07%	100,00%

Test for null hypothesis of normal distribution:
Chi-square(2) = 504,263 with p-value 0,00000

Рисунок 24 – Результаты теста на нормальность распределения остатков модели $wage = 6,74951 + 1,2509 \cdot educ$

4. ПОРЯДОК ВЫПОЛНЕНИЯ ЛАБОРАТОРНОЙ РАБОТЫ

1. Открыть набор данных File\Open Data\ Sample File (закладка Greene), Рисунок 25, в соответствии с номером варианта и общую текстовую информацию о нём.
2. Использовать 1МНК (OLS) для предварительной оценки параметров модели $y = \alpha_0 + \alpha_1 \cdot x + u$, оценить значимость параметров модели.
4. Определить, присутствует ли гетероскедастичность остатков модели.
5. В случае обнаружения гетероскедастичности провести оценку параметров модели методом ВМНК (WLS) двумя способами, оценить значимость параметров моделей.
6. Сравнить результаты, полученные с использованием методов 1МНК и ВМНК.



File	Summary	Вариант	ID#	x	y
greene5_1	U.S. macro data, 1950-2000	1		3	1
greene5_1	U.S. macro data, 1950-2000	2		5	1
greene7_8	Gasoline price and consumption	3		3	1
greene8_3	Solow (1957) production function data	4		4	2
greene11_3	Aggregate consumption and income data	5		2	1
greene12_1	Micro income and expenditure data	6		4	5
greene13_1	Grunfeld investment data	7		9	8
greene14_1	Airline costs data (six-firm panel)	8		5	3
greene18_2	GNP, money and price deflator	9		3	1
greene22_2	Fair data on extra-marital affairs	10		7	9

Рисунок 25 – Варианты заданий: название файла и номер варианта

5. СОДЕРЖАНИЕ ОТЧЕТА О ВЫПОЛНЕНИИ ЛАБОРАТОРНОЙ РАБОТЫ

- 1) Название и цель работы.
- 2) Постановка задачи.
- 3) Этапы выполнения задачи в Gretl.
- 4) Выводы.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Замков О.О. Взвешенный метод наименьших квадратов. Математические методы в экономике / О.О. Замков, А.В. Толстопятенко, Р.Н. Черемных. – М.: Дис, 1997. – 230 с.
2. Glesjer H. A New Test for Heteroscedasticity. Journ. Am. Stat. Assoc., 1969, vol. 64, pp. 316-323.
3. White H. A Heteroscedasticity-Consistent Matrix Estimator and a Direct Test for Heteroscedasticity. Econometrica, vol. 48 (1980), № 4, pp. 817-838.

Заказ № _____ от « _____ » _____ 2008. Тираж _____ экз.

Изд-во СевНТУ